



A Facial Feature-Based Binary Classification Framework for Deepfake Identification

Gaurav Aggarwal^{1*}, Gaurav Sharma¹, Ashutosh Bhatt¹, Pooja Joshi¹, Parul Saini²

¹ Department of Computer Science
Swami Rama Himalayan University
Dehradun, India
gaurav.aggarwalcoer@gmail.com
gauravsharma@srhu.edu.in
ashutoshbhatt15@gmail.com
poojajoshi.baloni@gmail.com

² School of Computing, DIT University
Dehradun, India parul.saini@dituniversity.edu.in

Abstract. Deepfake technology represents a growing threat within digital media, not only because of the risks of misinforming the public but also of privacy violations and security breaches. This research works on the binary classification of facial features to increase deepfake detection. Taking advantage of the inherent differences between real and fake faces, we presented an effective method using deep learning to identify minute changes in texture, lighting, and facial movements. Our technique is centered around the extraction of facial regions most likely to be altered (eyes, mouth, skin, etc.) through automation of image segmentation. The next step is to use multiple CNNs and dense layers to obtain a high-level feature representation and to categorize it into real or fake. Besides, our method utilizes data augmentation along with ensemble techniques to make it more resistant to changes in quality and other environmental factors. The accomplishment of the advanced performance model with minimum computing power use has proven the effectiveness of the proposed framework on the Kaggle dataset. In conclusion, our study points out binary classification of facial features could be a key method for identifying deepfakes in a manner that is accurate, reliable, and scalable. Thus, we are hopeful that it will lead to the practical application of media forensics and digital security.

Keywords: Deepfake image, CNN, Deepfake detection.

1 Introduction

In this digital era where pictures and videos dominate communication, the truthfulness of visual content has become a major aspect of credibility and trust.

*Corresponding author: gaurav.aggarwalcoer@gmail.com

© The Author(s) 2026

M. Singh et al. (eds.), *Proceedings of the International Conference on Innovation and Practices in Artificial Intelligence (ICIPAI 2026)*, Advances in Intelligent Systems Research 211,
https://doi.org/10.2991/978-94-6239-789-7_4

Mainly, facial images are a big deal because they influence so many sectors such as police, social networks, entertainment, and newspapers. On the other hand, visual data integrity is currently compromised by the increase of synthetic media, especially deepfake technology. Since these fakes use an advanced AI to generate very realistic fake face content, deepfakes have stirred a lot of concerns regarding misinformation, scams, and crimes. Hence, safeguarding the authenticity of facial characteristics in digital media is fundamental to keeping social order and public trust[27].

There are many and varied problems that face image manipulation using deepfake technology specifically. The difference from conventional photo and video editing methods is that deepfakes rely on neural networks to produce highly persuasive synthetic content so that even a human eye would have a hard time finding the signs. This has led to the exploitation of this weapon on a large scale, such as the production of fake news, identity theft, and other kinds of cybercrimes. Among the various methods in this field, one standout technique is studying the facial landmarks and how these change over time. Research has shown that visual clues of forgery are often found near certain facial features, which is why analyzing landmarks leads to a great method for detecting a wide range of forgeries[16][7]. As an example, one of the ways to identify face fakes is through the mesoscopic (intermediate-level) facial features, which are the target of MesoNet[1]. On the other hand, some approaches use 3D head pose estimation to spot discrepancies between facial landmarks and the overall orientation of the head [4]. Besides that, the shape ratios of different facial parts like the eyes, nose, and mouth can be calculated and monitored over time to detect unusual behavior in deepfake videos[5].

There has never been a greater need for efficient and scalable detection techniques given the seriousness of these issues. Deepfakes have been, if not impossible, very difficult to detect by the conventional detection methods such as manual inspection or metadata analysis, especially as the deepfake tools have progressed. AI- and machine learning-based methods have come to be one of the solutions. They attempt to differentiate original vs. manipulated images by detecting very small differences in features, motion, or texture [11].

The binary categorization of face features as a unique method for reliable deepfake identification is the main emphasis of this work. Many different approaches use the extensive technique of the analysis operation. This paper approaches detecting deepfakes based on a few special facial defects or mismatches that are introduced during the deepfake generation process. The purpose of the work is to enhance the accuracy and reliability of deepfake detection systems by focusing on fine-grained patterns in facial geometry, texture, and emotion[19].

It is true that industry has made many great achievements, but there are still numerous problems with the models of present time. Many of the detection techniques in use today, for example, suffer from inability to deal with all data sets, slow rate of processing, and being very easy to be hacked by the attackers. In addition, only very few models take into consideration the small facial features, which very often become the first clue of the alteration. Changes in physiology and behavior also

hold information that helps separate deepfakes from real ones. One well-known case of such detection is the analysis of eye-blinking patterns; since generative models in the past rarely generated natural blinking, tracking a whole physiological necessity like blinking became a way to check the truthfulness of a video[12]. Along the same lines, it has been recognized that out-of-sync lip movements and mouth actions are strong and widely applicable signs of detecting manipulated materials. Advanced methods even go as far as to pick up tiny physiological signs like the heart rate patterns through the technique of remote photoplethysmography (rPPG), which reveals the non-biological nature of artificial faces.

This research aims to accomplish two things. Firstly, it introduces a dual classifier tool for deepfake detection that relies on facial anomalies as the distinguishing feature. This technique is leveraging the latest deep learning framework for a more precise evaluation of the human face. Secondly, the article illustrates the operation of the proposed detection method through different sets of data, thereby providing evidence of its efficiency and scalability in the recognition of deepfakes. By addressing critical gaps in the domain, this publication makes a great contribution to the creation of efficient mechanisms for fighting the menace of deepfakes that is increasing day by day. The paper's objective is to improve the visual content verification sector through an in-depth analysis of deepfake detection issues, thereby ensuring the integrity and trustworthiness of digital media in the era of widespread fake content.

2 Literature Review

The growing use of deepfake technology has led to considerable research into the performance of various detection methods, mainly CNNs and advanced machine learning techniques. The research on deepfake binary facial image classification reveals that, in addition to custom CNN architectures, ResNet-50 and EfficientNet-B7 transfer learning frameworks are capable of accurately detecting the manipulated content. This paper highlights the essential requirement for strong detection techniques in order to counteract the increasing digital security threats due to deepfakes[10].

Ajoy et al. [1]teajoy2021deepfake design a deepfake detection system, based on CNNs, for finding pixel distortions and facial inconsistencies as major features of fake videos. The system contributes to social media platforms' efforts to reduce the dissemination of harmful fake videos and prevents misinformation and public unrest. Another research paper treats deepfake detection as a fine-grained classification task rather than the usual binary one. It helps improve cross-dataset and cross-manipulation robustness by learning subtle, generalizable features and using the background noise as a distraction. This essentially addresses the main weaknesses of current methods[18].

Jolly et al. [1][9] went deep to understand and automate the detection of facial expressions in video by working with Deepfake Face2Face, faceSwap, and neural texture. They chose a multilayered strategy, which helped them circumvent the issues

of data loss and compression. Apart from CNN-based models, one study explored the potential of generative adversarial networks (GANs) and data augmentation to create datasets. The experiments with pretrained models such as VGG16 and ResNet50 demonstrated significant increases in accuracy, more so with the ensemble methods that integrate these models to reach accuracies of up to 98.79% on standard datasets[23].

Altaei et al. [1]tealtaei2022detection have implemented CNN for the detection of fake images. Their method consists of image preprocessing, gamma correction, and Canny filter extraction. The authors apply two detection procedures: CNN combined with Principal Component Analysis (PCA) and CNN without PCA, respectively. Although the results reveal a good level of accuracy, the CNNonly method yields the highest accuracy. Lastly, an original methodology incorporates scalable CNNs, like EfficientNet, with hierarchical vision transformers (ViT), particularly the shifted window transformer. This combination of CNNs and transformers, exemplified by the hybrid architecture, has resulted in an excellent accuracy of 98.04%, which is quite promising for deepfake detection[3]. The rapid increase in deepfake technology, which has been spurred by the great progress made in Generative Adversarial Networks (GANs) and Variational Autoencoders (VAEs), has raised a serious issue, the need for strong systems to protect the digital world from being compromised. Convolutional Neural Networks (CNNs) have become the go-to means for detecting such synthetic changes mainly because they can independently learn and extract characteristic spatial features from images[6]. Initially, the researchers were looking for identifying mesoscopic properties, features that are at a medium scale and that generative models usually fail to replicate accurately. Afchar et al. (2018) developed MesoNet, a small CNN layout that used Inception-like blocks to efficiently detect facial forgeries; however, its capacity was later shown to be rather weak against advanced, high-resolution deepfakes[1].

After that, generative models changed, and in response, computer vision experts looked for deeper and more complicated CNN structures that would enable them to detect smaller and more subtle artifacts. XceptionNet, a neural network architecture that applies depthwise separable convolutions, was set as the standard in the domain following the study by Rossler et al. (2019) that revealed that it outperformed competitors on the FaceForensics++ dataset, reaching almost perfect accuracy under lab conditions[21]. That said, there's still the issue of generalizing to other contexts, since the models that were fit to certain datasets don't often work well when they see "in the wild" deepfakes. For this reason, Tan and Le (2019) came up with EfficientNet, which uses a compound scaling approach to justify increasing network depth, width, and resolution, thus coming up with a more reliable framework for feature extraction with fewer parameters[26]. EfficientNet, B4, has been the one that is most commonly used due to its effectiveness in detecting minute deviations in facial textures and lighting levels[2]. Nowadays new papers are analyzing a combination of transfer learning and pretrained models to boost detection accuracy. Say VGGFace and ResNet architectures are examples where their face-related features can be helpful in finding deepfakes[20]. Shad et al. (2021) compared eight various CNN architectures, including

DenseNet and ResNet; by performing their experiments, they have confirmed that, although deep networks have great accuracy, they also have vulnerabilities, for example, to adversarial attacks or compression artifacts[22]. Therefore, this flaw has made these developers come up with even more focused networks such as DeepGuardNet (2025), which is planned to be capable of grabbing the characteristics/features of images in both the space and depth directions so as to be more unbreakable against fake art/content [24].

The field is now moving towards hybrid models that exploit the local spatial processing power of CNNs but, at the same time, take advantage of other deep learning paradigms. RNNs and LSTM units are often used together with CNN backbones, for example, in the framework by Masood et al. (2023) [1] to study temporal discrepancies among video frames. Besides, Vision Transformers (ViTs) have offered a completely new technique of deepfake detection. Wodajo and Atnafu (2021) came up with a blend of CNN and Transformer models in which a CNN is used to extract local features and a Transformer is used to capture global dependencies, thus effectively dealing with the major drawback of CNNs in understanding long-distance relations of an image[29].

Datasets such as Celeb, DF, and the Deepfake Detection Challenge (DFDC) were instrumental in measuring the performance of these models. Li et al. (2020) remarked that although many CNN-based detectors still perform quite well on the first-generation deepfakes, their performance drops significantly when they are tested with high-quality datasets like Celeb-DF, which have clearly realistic blending and hardly any visual artifacts[13]. This has therefore opened the door to the investigation of explainable AI (XAI) for deepfake detection, where the intention is to make it possible to see the very parts of the image that a CNN predicts as "fake" and so help to a greater extent the transparency as well as the trustworthiness of the detection process[17].

Nevertheless, some difficulties remain in tackling synthetic media despite these technological improvements. Malik et al. (2022) argued that the "cat and mouse" chase between generators and detectors would go on since new GAN designs are made without fail to circumvent existing CNN-based forensic methods[15]. Besides, the situation is quite complex when it comes to the influence of environmental factors such as a low level of illumination, motion blur, and high compression, which pose a serious challenge for the practical use of such systems[14]. As a result, the research in this domain may primarily revolve around self-supervised learning as well as multimodal detection, i.e., integrating visual, audio, and metadata analysis for the development of more robust and broadly capable identification methods[25].

3 Proposed Deepfake Detection

In this part, a method is proposed for identifying deepfake images. One of the main aids in detecting deepfakes, especially in a two-class problem using Convolutional Neural Networks (CNNs), is a dataset comprising approximately 190,000 real and fake

facial images. A summary of all dataset details can be found in Table 1. This collection of real and manipulated images of a dataset is not only well-balanced but also diverse, showing variations in lighting, facial expressions, and methods of tampering. Several characteristics are necessary for training CNNs so that they perform well when generalizing to real-world conditions. Since CNNs excel at learning features at various levels, they can detect very small inconsistencies caused by deepfake generation, for example, in texture, lighting, and facial alignment. Using this dataset as a source for training and evaluating deepfake detection methods, researchers can come up with strong approaches for confronting deepfake issues. Such inventions not only combat misinformation dissemination but also reinforce the digital media authentication systems. The dataset’s operationalization is depicted in Figure 1.

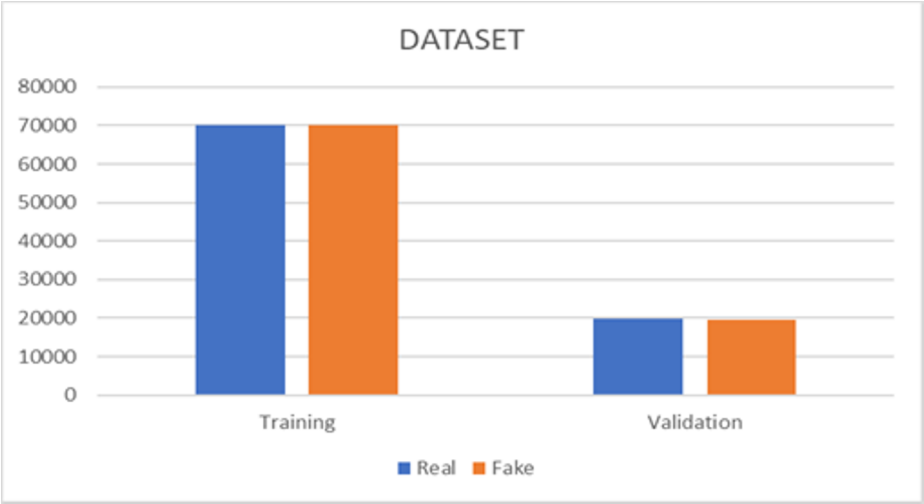


Fig.1. Implementation details of the Dataset

Table 1. DATASET USED IN PROPOSED MODEL

Dataset	Images	Resolution
190K Fake and Real Faces images	190000	256*256

The design of the proposed architecture is elaborated in Figure 2. Our approach is based on CNN architecture, which is targeted specifically for detecting deepfake images. First of all, we wanted to create a model that is able to extract

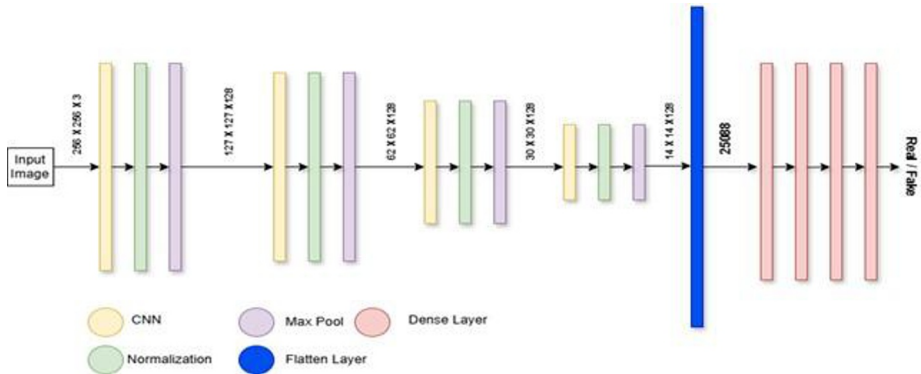


Fig.2. Architecture of Proposed Model for Binary classification of Deepfake Image

spatial features from images and then, based on these features, automatically categorize images as real or fake. The model accepts an RGB image as input, and the image is of a fixed size of 256x256 pixels. The model's first layers consist of Conv2D layers in a sequence, with 128 filters in each one and 3x3 kernels. In general, these layers identify the lowest-level features such as edges, textures, and very simple patterns.

A batch normalization layer is added after every convolutional operation, normalizing the activations and allowing the training to be faster. To further reduce computational complexity, we have also employed Max Pooling layers. These layers reduce the resolution of the feature maps by selecting the most important values from the local neighborhoods. The model has four convolutional blocks, each one progressively reducing the spatial dimensions of the feature maps. The input size of 256x256 is effectively being converted into 14x14 through the feature maps. By performing this reduction, the model is able to focus on high-level spatial features while it remains computationally efficient. Then, the output of the convolutional blocks is fed into a Flatten layer, which changes the 3D feature map into a 1D vector of length 25088.

This transformation makes the data compatible with fully connected layers, which perform the classification. The 1D vector features are processed through several dense layers, which have 128 neurons each. Dropout layers are placed between the dense layers for the purpose of preventing the model from overfitting and improving its generalization capability. In training, these layers deactivate randomly a certain fraction of neurons, which ensures that the model doesn't depend too much on certain features.

The last dense layer only has one neuron, which produces a binary value showing if the input image is real or fake. The model has about 3.8 million parameters in total, out of which 3.7 million can be trained. This model manages to strike a balance between complexity and efficiency and is therefore capable of detecting deepfake images with high accuracy while still being computationally reasonable. By effectively

integrating convolutional, pooling, and dense layers, the model is able to capture both local and global features, thus enhancing its robustness to different deep fake styles. As shown in Table 2, the proposed detailed model architecture consists of multiple layers along with trainable parameters designed for effective deepfake image classification. The parameter settings for the CNN-based classifier included Adam as the optimizer, an initial learning rate of 0.001, a loss of binary crossentropy, and a batch size of 32 and trained the CNN classifier for 10 epochs. The data set was partitioned into training and validation data by an 80:20 split. The training data was used for training the model, while the validation data was used for comparison and to avoid overfitting. The optimizer parameters were updated according to validation accuracy, which was used as the primary evaluation index.

Table 2. Summary of the Architecture of Proposed Model for binary classification of Deepfake Images

Layer (Type)	Output Shape	Param #
Conv2D	(None, 254, 254, 128)	3,584
BatchNormalization	(None, 254, 254, 128)	512
MaxPooling2D	(None, 127, 127, 128)	0
Conv2D	(None, 125, 125, 128)	147,584
BatchNormalization	(None, 125, 125, 128)	512
MaxPooling2D	(None, 62, 62, 128)	0
Conv2D	(None, 60, 60, 128)	147,584
BatchNormalization	(None, 60, 60, 128)	512
MaxPooling2D	(None, 30, 30, 128)	0
Conv2D	(None, 28, 28, 128)	147,584
BatchNormalization	(None, 28, 28, 128)	512
MaxPooling2D	(None, 14, 14, 128)	0
Flatten	(None, 25088)	0
Dense	(None, 128)	3,211,392
Dropout	(None, 128)	0
Dense	(None, 128)	16,512
Dropout	(None, 128)	0
Dense	(None, 128)	16,512
Dropout	(None, 128)	0
Dense	(None, 128)	16,512
Dropout	(None, 128)	0
Dense (Output)	(None, 4)	129
Total Params	3,709,441	
Trainable Params	3,708,417	
Non-trainable	1,024	

4 Experiments and Results

The proposed model is highly effective on a sample of 190,000 genuine and fake face photos, as evidenced by the results in Tables 3 and 4. It achieved a training accuracy of 99.23%, indicating its ability to effectively extract knowledge from the training data. Its 94.5% F1 measure guarantees a reliable classification capability by duly representing the balance between precision and recall. performance.

Besides, a precision level of 95% that measures the proportion of true positive

Table 3. Comparison of the proposed model's performance with other approaches on the dataset.

Model Name	Test Accuracy (%)
ResNet50[8]	53.43
FaceNet [11]FaceNet[28]	94.51
Conv2D CNN[4]	95.85
Proposed Model	99.23

Table 4. PERFORMANCE EVALUATION OF THE PROPOSED MODEL

Model	Train Accuracy	F1 Score	Precision	Recall
Proposed Model	99.23	94.5	95	94

cases in all positive cases recognized by the model shows that the model is doing a great job in recognizing fake faces with a low level of false alarm.

The 94% recall score of the model further substantiates its trustworthiness in distinguishing between real and fake faces, as it measures how well the model is identifying the positive cases. These results showcase the efficiency and robustness of the model in addressing this challenging categorization problem. From Table 3 and Table 4, the following observations have been made:

- A deepfake detection model that has been proposed outperforms other existing ones greatly, both in terms of test accuracy (99.23%) and robustness. %.
- Having a train accuracy of 99.23% implies that the network has successfully learned the main features.
- An F1 score of 94.5% indicates that both precision and recall have been well balanced.
- High precision of 95% will result in fewer false positives, while recall of 94% will help in identifying fake images efficiently.

The binary class, fake and real, classification confusion matrix has been depicted in Figure 3. A confusion matrix of size 2x2 (for binary classification) associated

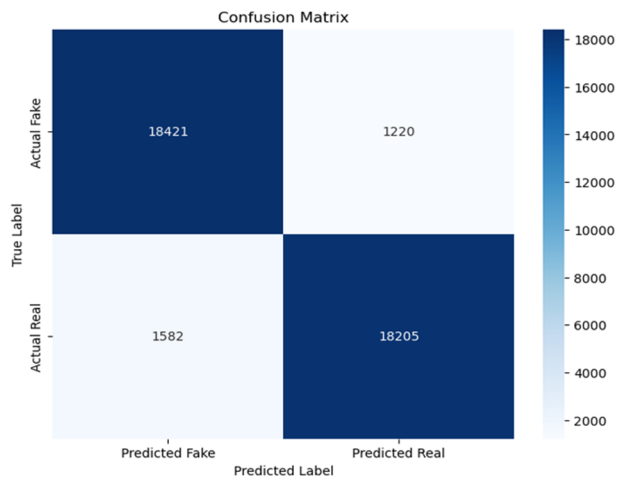


Fig.3. Confusion Matrix

with a classifier displays expected and actual classifications, where 2 is the number of unique classes. The confusion matrix for the proposed model on the 190k Real and Fake Faces dataset illustrates that the true positives and true negatives are on the diagonal elements.

5 Conclusion

The new model we offer is a holistic way of detecting deep fakes by utilizing CNNs’ capability to recognize very complex patterns and irregularities. The CNN-based model we suggest outperforms other models in detecting deepfake images, with a high accuracy rate of 99.23%. It shows great progress compared to ResNet50, FaceNet, and other Conv2D CNN methods. The model’s outstanding ability to extract and analyze features mainly for uncovering deep fake elements highlights the importance of designing dedicated architecture in this domain. Incorporating batch normalization, dropout, and several dense layers does not merely enhance the model accuracy but also adds a layer of overfitting prevention, thereby making it really apt for the re-worlding scenario of detecting altered media.

References

1. Afchar, D., Nozick, V., Yamagishi, J., Echizen, I.: Mesonet: a compact facial videoforgery detection network. In: 2018 IEEE international workshop on information forensics and security (WIFS). pp. 1–7. IEEE (2018)
2. Agnihotri, A.: DeepFake Detection using Deep Neural Networks. Ph.D. thesis,

Dublin, National College of Ireland (2021)

3. Ajoy, A., Mahindrakar, C.U., Gowrish, D., et al.: Deepfake detection using a framebased approach involving cnn. In: 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA). pp. 1329–1333. IEEE (2021)
4. Altaei, M.S.M., et al.: A detection of deep fake in face images using deep learning. Wasit Journal of Computer and Mathematics Science **1**(4), 60–71 (2022)
5. Carter, B., Dilla, N., Callahan, M., Ambala, A.: Deepfake detection via facialfeature extraction and modeling. arXiv preprint arXiv:2507.18815 (2025)
6. Edwards, P., Nebel, J.C., Greenhill, D., Liang, X.: A review of deepfake techniques: Architecture, detection, and datasets. IEEE Access **12**, 154718–154742 (2024)
7. Gao, Q., Zhang, B., Wu, J., Luo, W., Teng, Z., Fan, J.: Leveraging facial landmarksimproves generalization ability for deepfake detection. Pattern Recognition **164**, 111528 (2025)
8. Hettiarachchi, S.: Analysis of different face detection andrecognition models forandroid (2021)
9. Jolly, V., Telrandhe, M., Kasat, A., Shitole, A., Gawande, K.: Cnn based deeplearning model for deepfake detection. In: 2022 2nd Asian conference on innovation in technology (ASIANCON). pp. 1–5. IEEE (2022)
10. Kalemullah, A.M., Prakash, P., Sakthivel, V.: Deepfake classification for humanfaces using custom cnn. In: 2024 7th International Conference on Circuit Power and Computing Technologies (ICCPCT). vol. 1, pp. 744–750. IEEE (2024)
11. Khalil, H.A., Maged, S.A.: Deepfakes creation and detection using deep learning.In: 2021 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC). pp. 1–4. IEEE (2021)
12. Li, Y., Chang, M.C., Lyu, S.: In icu oculi: Exposing ai created fake videos by detecting eye blinking. In: 2018 IEEE International workshop on information forensics and security (WIFS). pp. 1–7. IEEE (2018)
13. Li, Y., Yang, X., Sun, P., Qi, H., Lyu, S.: Celeb-df: A large-scale challenging datasetfor deepfake forensics. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 3207–3216 (2020)
14. Lipianina-Honcharenko, K., Telka, M., Melnyk, N.: Comparison of resnet, efficientnet, and xception architectures for deepfake detection. In: AdvAIT. pp. 26–34 (2024)
15. Malik, A., Kuribayashi, M., Abdullahi, S.M., Khan, A.N.: Deepfake detection forhuman face images and videos: A survey. IEEE Access **10**, 18757–18775 (2022)
16. Masi, I., Killekar, A., Mascarenhas, R.M., Gurudatt, S.P., AbdAlmageed, W.: Twobranch recurrent network for isolating deepfakes in videos. In: European conference on computer vision. pp. 667–684. Springer (2020)
17. Masood, M., Nawaz, M., Malik, K.M., Javed, A., Irtaza, A., Malik, H.: Deepfakesgeneration and detection: state-of-the-art, open challenges, countermeasures, and way forward: Deepfakes generation and detection: state-of-the-art, open challenges, countermeasures, and way forward. Applied intelligence **53**(4), 3974–4026 (2023)
18. Nadimpalli, A.V., Rattani, A.: Facial forgery-based deepfake detection using finegrained features. In: 2023 International Conference on Machine Learning and Applications (ICMLA). pp. 2174–2181. IEEE (2023)
19. Nguyen, T.T., Nguyen, Q.V.H., Nguyen, D.T., Nguyen, D.T., Huynh-The, T., Nahavandi, S., Nguyen, T.T., Pham, Q.V., Nguyen, C.M.: Deep learning for deepfakes creation and detection: A survey. Computer Vision and Image Understanding **223**, 103525 (2022)

20. Raza, A., Munir, K., Almutairi, M.: A novel deep learning approach for deepfake image detection. *Applied Sciences* **12**(19), 9820 (2022)
21. Rossler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., Nießner, M.: Faceforensics++: Learning to detect manipulated facial images. In: *Proceedings of the IEEE/CVF international conference on computer vision*. pp. 1–11 (2019)
22. Shad, H.S., Rizvee, M.M., Roza, N.T., Hoq, S.A., Monirujjaman Khan, M., Singh, A., Zaguia, A., Bourouis, S.: [retracted] comparative analysis of deepfake image detection method using convolutional neural network. *Computational intelligence and neuroscience* **2021**(1), 3111676 (2021)
23. Sharma, J., Sharma, S., Kumar, V., Hussein, H.S., Alshazly, H.: Deepfakes classification of faces using convolutional neural networks. *Traitement du Signal* **39**(3) (2022)
24. Simon, P., et al.: Deepguardnet: A novel cnn architecture for deepfake image detection. *Procedia Computer Science* **258**, 811–818 (2025)
25. Sunil, R., Mer, P., Diwan, A., Mahadeva, R., Sharma, A.: Exploring autonomous methods for deepfake detection: A detailed survey on techniques and evaluation. *Heliyon* **11**(3) (2025)
26. Tan, M., Le, Q.: Efficientnet: Rethinking model scaling for convolutional neural networks. In: *International conference on machine learning*. pp. 6105–6114. PMLR (2019)
27. Tolosana, R., Vera-Rodriguez, R., Fierrez, J., Morales, A., Ortega-Garcia, J.: Deepfakes and beyond: A survey of face manipulation and fake detection. *Information fusion* **64**, 131–148 (2020)
28. Wen, L., Xu, D.: Face image manipulation detection. In: *IOP conference series: materials science and engineering*. vol. 533, p. 012054. IOP Publishing (2019)
29. Wodajo, D., Atnafu, S.: Deepfake video detection using convolutional vision transformer. *arXiv preprint arXiv:2102.11126* (2021)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

